

```

GTGACTGCTATCACCCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTGCACT
1 -----+-----+-----+-----+-----+ 60
  M T A I T L G G L L L K G I I T L V S T
CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAAACCGCGGCGAAATC
61 -----+-----+-----+-----+-----+ 120
  H T S S T L P A F K K I L V A N R G E I
GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGTGCAGCCACGCTAGCTATTTACCCC
121 -----+-----+-----+-----+-----+ 180
  A V R A F R A A L E T G A A T V A I Y P
CGTGAAGATCGGGGATCATTCACCGCTCTTTTGCTTCTGAAGCTGCCGATTTGGTACT
181 -----+-----+-----+-----+-----+ 240
  R E D R G S F H R S F A S E A V R I G T
GAAGGCTCACCACTCAAGCGTACCTGGACATCGATGAAATTTATCGGTGCAGCTAAAAAA
241 -----+-----+-----+-----+-----+ 300
  E G S P V K A Y L D I D E I I G A A K K
GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAATGCCAGCTTGCC
301 -----+-----+-----+-----+-----+ 360
  V K A D A I Y P G Y G F L S E N A Q L A
CGCGAGTGC CGGAAAAACGGCATTACTTTTATTGGCCCAACCCAGAGGTCTCTGATCTC
361 -----+-----+-----+-----+-----+ 420
  R E C A E N G I T F I G P T P E V L D L
ACCGGTGATAAGTCTCGTGCCTAACC CGCGCAAGAAGGCTGGTCTGCCAGTTTGGCG
421 -----+-----+-----+-----+-----+ 480
  T G D K S R A V T A A K K A G L P V L A
GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGCCAGACTTAC
481 -----+-----+-----+-----+-----+ 540
  E S T P S K N I D D I V K S A E G Q T Y
CCCATCTTTGTAAAGGCAGTTGCCGTTGGTGGCGGACGCGGTATGCGCTTTGTTCTTCA
541 -----+-----+-----+-----+-----+ 600
  P I F V K A V A G G G G R G M R F V S S
CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC
601 -----+-----+-----+-----+-----+ 660
  P D E L R K L A T E A S R E A E A A F G
GACGTTCCGTATATGTGCAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC
661 -----+-----+-----+-----+-----+ 720
  D G S V Y V E R A V I N P Q H I E V Q I

```

FIG. 1A

```

CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT
721 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 780
    L G D R T G E V V H L Y E R D C S L Q R
CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT
781 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 840
    R H Q K V V E I A P A Q H L D P E L R D
CGCATTGTGCGGATGCGATAAAGTTCTGCCGCTCCATTGGTTACACGGGCGCGGGAACC
841 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
    R I C A D A V K F C R S I G Y Q G A G T
GTGGAATTCTTGGTCGATGAAAAGGGCAACCAGTTTTCATCGAAATGAACCCACGTATC
901 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 960
    V E F L V D E K G N H V F I E M N P R I
CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG
961 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020
    Q V E H T V T E E V T E V D L V K A Q M
CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCCAAGATAGATCAAGACC
1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
    R L A A G A T L K E L G L T Q D K I K T
CACGGTGCAGCACTGCAGTGCCGCATCACCCAGGAAGATCCAAACAACGGCTTCCGCCCA
1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140
    H G A A L Q C R - I T T E D P N N G F R P
GATACCGGAAGTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT
1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
    D T G T I T A Y R S P G G A G V R L D G
GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC
1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260
    A A Q L G G E I T A H F D S M L V K M T
TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGGTTGGCTGAGTTC
1261 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1320
    C R G S D F E T A V A R A Q R A L A E F
ACCGTGTCTGGTGTGCAACCAACATTTGGTTTCTTGCGTGCGTGTCTGCGGGAAGAGGAC
1321 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1380
    T V S G V A T N I G F L R A L L R E E D
TTCACTTCCAAGCGCATCGCCACCGGATTATCGGCGATCACCCACACCTCCTTCAGGCT
1381 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1440
    F T S K R I A T G F I G D H P H L L Q A

```

FIG. 1B

```

1441 CCACCTGCGGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC
-----+-----+-----+-----+-----+-----+-----+-----+
      P P A D D E Q G R I L D Y L A D V T V N
1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTGCAGCACCAATCGATAAGCTGCCAACATC
-----+-----+-----+-----+-----+-----+-----+-----+
      K P H G V R P K D V A A P I D K L P N I
1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCAGCCGCG
-----+-----+-----+-----+-----+-----+-----+-----+
      K D L P L P R G S R D R L K Q L G P A A
1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTTGATACCACCTTCCGCGAT
-----+-----+-----+-----+-----+-----+-----+-----+
      F A R D L R E Q D A L A V T D T T F R D
1681 GCACACAGTCTTTGTGTGCGACCCGAGTCCGCTCATTTCGCACTGAAGCCTGCGGCAGAG
-----+-----+-----+-----+-----+-----+-----+-----+
      A H Q S L L A T R V R S F A L K P A A E
1741 GCGCTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGCGCTGGGGCGGCGCACTTAC
-----+-----+-----+-----+-----+-----+-----+-----+
      A V A K L T P E L L S V E A W G G A T Y
1801 GATGTGGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG
-----+-----+-----+-----+-----+-----+-----+-----+
      D V A M R F L F E D P W D R L D E L R E
1861 GCGATGCCGAATGTAACATTCAGATGCTGCTTCGCGGCGCAACACCGTGGGATACACC
-----+-----+-----+-----+-----+-----+-----+-----+
      A M P N V N I Q M L L R G R N T V G Y T
1921 CCGTACCCAGACTCCGTCTGCCGCGGTTTGTTAAGGAAGCTGCCAGCTCCGCGCTGGAC
-----+-----+-----+-----+-----+-----+-----+-----+
      P Y P D S V C R A F V K E A A S S G V D
1981 ATCTTCGCAATCTTCGACGCGCTTAACGACGCTTCCAGATGCTCCAGCAATCGACGA
-----+-----+-----+-----+-----+-----+-----+-----+
      I F R I F D A L N D V S Q M R P A I D A
2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATCTGTTGATCTCTCT
-----+-----+-----+-----+-----+-----+-----+-----+
      V L E T N T A V A E V A M A Y S G D L S
2101 GATCCAAATGAAAGCTCTACACCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC
-----+-----+-----+-----+-----+-----+-----+-----+
      D P N E K L Y T L D Y Y L K M A E E I V
2161 AAGTCTGGCGCTCACATTTCTGCCATTAAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG
-----+-----+-----+-----+-----+-----+-----+-----+
      K S G A H I L A I K D M A G L L R P A A
2220

```

FIG. 1C